

WEB BROWSER EXTENSION AND METHOD FOR PROCESSING
DATA CONTENT OF WEB PAGES

5

Field of Invention

10 The present invention relates to extending Web Browser capabilities and in particular to assisting a user of a Web Browser with analysis and interpretation of the data content of displayed Web pages.

Background

15 The majority of the content of Web pages consists of text, images, and other media together with the HyperText Markup Language (HTML) or other markup language codes which describe how the media content should be displayed by a Web Browser. This content is typically static and cannot be manipulated by the user.

20

Some Web pages also contain embedded active components that support manipulation of content on a specific page within constraints defined by the page designer. Examples of such components which are integrated within individual Web pages and limited to those pages include Java applets and ActiveX controls. Many controls of this kind impose a per-page download penalty that can impact the performance of the Web page. Another type of embedded active component is plug-in applications for Netscape's Navigator Web Browser, which are each associated with a specific type of non-HTML data content and allow users to 'view' or 25 experience that content. Example plug-ins are Adobe's Acrobat document viewer and Macromedia's Shockwave interactive animation and sound player.

30

Other functions for working with Web pages have been implemented recently as Browser extensions. These have been used, for example, to reserve user interface space within the Browser for displaying links to other Web pages or links to other applications. The Alexa Browser extension performs a database lookup using the URL of the current Web page and displays a variety of content related to this URL. Typically, linked applications are stand-alone and have no particular relationship 35 to the Browser content once launched and include no mechanism for communicating back to the Browser. In other cases, the browser extensions have enabled simple manipulation of Web page content from context menus, such as applying a background colour to a selected block of text to 40 highlight it, or listing and displaying the images or the links in the current Web page. In all of these cases, no analysis or interpretation of 45 the semantics of the Web page data content is performed. Furthermore, to

the extent that a user is able to select a portion of data content to be manipulated, that portion of content is typically manipulated as a block without separation of individual data items within the selected portion of content.

5

AltaVista Babel Fish is a Web page content translation tool. The user links to the Babel Fish Web site by clicking on the Babel Fish icon in the toolbar of Microsoft's Internet Explorer 5. The user chooses a language and specifies a URL or selected text via an intermediate screen (typing in data or using standard copy and paste functions) and then the translation is performed on AltaVista's relevant server computer. While the Babel Fish tool adds to the capabilities of a Web Browser and provides a method for manipulation of Web page content, the only new component implemented on the Web client computer is a mechanism for navigating to the Babel Fish site. Note that Babel Fish does not perform the translation processing on the data processing apparatus which is running the Web Browser (i.e. on the Web client) but on a remote server. After the tool has been used, the user has to navigate back to the original Web page. Functions as computationally complex as translation require more processing power than is typically available on a Web client device and considerable storage space is required to hold dictionaries in multiple languages, and so there are significant technical reasons against implementing Babel Fish translation on a Web client device.

25

Thus, existing Web Browsers, embedded components, Browser extensions and other associated applications have significantly limited capabilities for manipulation of Web page data content and in particular for analysis and interpretation of semantic data content and for manipulation of individual data items of a set of items selected from a Web page by a user.

30

Summary of Invention

35

In a first aspect, the invention provides a set of components for cooperating with a Web Browser to process the semantic data content of Web pages, wherein the Web Browser implements functions to display Web page contents on a display apparatus, and wherein the set of components are operable to run on a Web client data processing apparatus together with the Web Browser, the set of components including: a data content transfer component, responsive to user selection of a desired semantic processing operation and user selection of a set of data items within a displayed Web page, for extracting the set of data items from the Web page; and a set of data content processing components, each for processing individual data items of a set of data items extracted from any one of multiple Web pages, for performing at the Web client data processing apparatus the selected processing operation on the individual

40

45

items of the selected set, to generate a result which is dependent on the semantics of the selected set of data items; wherein the data content transfer component includes means for providing the generated result to the Web Browser for presentation to the user.

5

The invention extends a Web Browser by providing the capability to process semantic data content of Web pages - this means performing processing operations on actual data values where the values of the data (e.g. number values or the meaning of text words) are relevant to the result of the operation. The invention helps users to analyze and interpret text or numeric content displayed by the Browser. This goes beyond the capabilities of existing Browsers which typically only include functions for accessing Web pages and parsing and rendering Web page content for display to a user. It also goes beyond the capabilities of known Web Browser extensions which typically only provide links to related Web pages or other Web pages of interest to the user, or which provide links to other applications without providing any ability to process the semantic data content of a Web page or to return the results of the processing to the Browser.

10

15 The set of processing components are each generically applicable to multiple Web pages (preferably working with any HTML-based content on any Web page, or any markup language) and so are distinguished from Java applets, ActiveX controls and other known active components which are embedded within a single, specific Web page and perform operations on the data within a specific area of that Web page only within constraints defined by the Web page designer. This generic applicability is achieved by implementing the data content transfer component at the Browser level rather than at the page level.

20

25 The set of processing components are 'client-side tools', running on the data processing apparatus on which the Web Browser runs. This distinguishes the components of the present invention from a tool which requires a processing request and the client data to be sent across an Internet connection to a remote server which processes the data remotely and then returns the results of the processing to the client system. Preferred embodiments of the invention therefore avoid the delay and scope for failures inherent in network transmission and no per-page download overhead applies.

30

35 Although the processing operation is performed locally, some implementations of the present invention may include a retrieval of information from a remote computer for use in the processing. One example is retrieval of current taxation rates from a remote database for use in calculating tax payable on net prices listed in a displayed Web page.

40

The components of the invention preferably provide the results of their processing to the Web Browser to display as modified HTML content (or XML, WML or other content-description or 'markup' language of the original content), the result being represented in a pop up window or via some form of annotation of the Web page. Alternatively, the result may be displayed elsewhere in the Browser's window outside the area reserved for display of the Web page. Note that these are examples of the processing results being displayed so as to modify the information displayed by the Browser while still displaying the Web page which was displayed when the processing component and Web page data contents were selected.

A Web Browser suitable for implementation of the present invention may comprise a tool for responding to user interactions by sending to a Web server HTTP or other requests for Web pages and for presenting retrieved pages to the user (for example, providing functions for submitting requests to an Internet search engine located elsewhere in the network and for parsing and rendering received HTML). In the present application, the term 'Web Browser' is also intended to include any markup language viewer component which implements functions for displaying markup language content to a user of a data processing apparatus, and which may or may not include other functions implemented by conventional Web Browsers. The term 'Web client apparatus' should be interpreted accordingly, and may include communications-enabled PDAs, telephones with data processing capability, in-car devices, television sets and other devices. A Web client apparatus in this context may also be a device providing non-visual interfaces for accessing Web page content such as using a voice-enabled Browser to access VoiceXML content. Additionally, 'Web Browsers' suitable for implementing the invention may include significant additional functionality, such as data manipulation functions.

The present invention preferably provides a generic enablement of the Browser to work with any one of a plurality of data processing tools which are able to exchange data with the Browser-enabling code, and is therefore open-ended. That is, additional processing components can be added to this infrastructure after the enabling code has been installed on a client device. The individual processing components can typically be implemented as very small code components, suitable for storing and running on a client device with limited storage.

The present invention also provides significant advantages over the invention described in US patent No. 5,896,491, which is incorporated herein by reference. US 5,896,491 describes a set of generic, client-side tools for performing processing operations in response to user selection of an operation and selection of a set of an application's data items and applying emphasis such as highlighter colours to particular data values.

to aid interpretation and analysis. A limitation of the invention of US 5,896,491 was that it required each application program to be engineered to support interoperability with the generic tools. By contrast, the set of processing components of a Web Browser according to the preferred embodiment of the present invention can access and manipulate data content in any typical HTML-based Web page, including applications built using HTML, (or another markup language).

According to a preferred embodiment of the present invention, numbers and strings can be extracted from unstructured Web content and the extracted data is then passed in a structured form to a processing component. Thus, processing components which are adapted to handle structured, numerical, tabular data are enabled to work with unstructured data in Web pages.

In a second aspect of the invention, there is provided a method for processing the semantic data content of Web pages displayed on a display apparatus via a Web Browser, the method including the following steps performed on a Web client data processing apparatus: in response to user selection of a desired semantic processing operation, and user selection of a set of data items within a displayed Web page, extracting the set of data items from the Web page; and performing the selected processing operation on the selected set of data items to generate a result which is dependent on the semantics of the selected set of data items, using a data content processing component which is capable of processing the data content of any one of multiple Web pages; providing the generated result to the Web Browser for presentation to the user.

In a third aspect, the invention provides a software component for cooperating with a Web Browser, the software component controlling the operation of a data processing apparatus on which it runs to perform the following steps in response to user selection of a desired data processing operation and user selection of a portion of a Web page displayed via the Browser: extracting the selected portion of the markup language Web page content from Web page; separating individual data items from markup tags within the selected portion; passing the separated data items to a data processing component to perform a processing operation on the individual data items; and in response to receipt by said software component of a result of said processing operation, generating a modified version of the portion of the markup language Web page content; and providing said modified portion to the Brower to replace the extracted portion of the displayed Web page.

This extraction of a selected portion of a Web page's markup language content, followed by separation of individual data items for processing, and then generation of a modified portion of markup language

5 to replace the extracted portion is distinguished from the prior art manipulation of blocks of text which does not involve an identification of the separate data items. In addition to a single word or number, an 'individual data item' in this context may comprise a set of tokens which form a single logical unit, such as a name combining a forename token and a surname token. The majority of Web page content is currently written in HTML, but the invention can be applied to any markup language.

10 This aspect of the invention of extraction of individual data values from a selected portion of a Web page's markup content, sending to a processing component, and generating modified markup content in response to receipt of a processing result is applicable to many different types of processing operation. Such operations may be, for example, a table lookup operation to obtain information associated with a data item, a chart generation operation, or a mathematical calculation.

20 The present invention may be implemented in a computer program product comprising computer readable program code recorded on a computer readable recording medium. Alternatively, the program code may be provided on a Web server computer in a form downloadable to Web client data processing systems. The invention can be integrated within a Web Browser as well as being provided as a component or set of components for interoperating with a Browser.

25 The invention according to a further aspect comprises a data processing system including a set of software components for cooperating with a Web Browser to enable processing of individual data items of a set of data items extracted from the data content of Web pages.

30 Brief Description of Drawings

35 Embodiments of the present invention will now be described in more detail, by way of example, with reference to the accompanying drawings in which:

Figure 1 is a schematic representation of the main software components, interface objects and their relationships, according to an embodiment of the invention; and

40 Figure 2 is a screen shot image showing an example implementation of the present invention in which the user interface includes an explorer bar and palette for user interaction, and in which the selected 'News Reader' operation has identified particular categories of company and generated emphasis results (background colours) which have then been applied to the displayed company names, and generated annotation results which have been displayed on a Browser status line;

Detailed Description of Preferred Embodiment

A first embodiment of the invention uses a new type of Browser extension to Microsoft's Internet Explorer 5 Web Browser to provide a mechanism linking the selection model of the Browser to a plurality of data content processing components known as SmartMarkers. The invention may be implemented in any data processing apparatus which is suitable for use as a Web client. Such a device is capable of running a Web Browser and typically includes a processor, memory for storing data and program code (which may include an internal hard disk device or an external storage device), a display device connected to receive data via a display controller, one or more input devices (typically a mouse and a keyboard) and an input controller, and appropriate internal and external communications resources. The SmartMarker data content processing components are preferably implemented as Java(TM) program components. SmartMarkers are general purpose end-user tools for analysing and interpreting text or numeric content displayed by a Browser or SmartMarker-enabled application program. The design of SmartMarkers and a number of specific examples is described in UK patent application No. GB 9816098.9 (corresponding to PCT/GB99/02433), which is incorporated herein by reference and which includes a description of annotating application data values as a way of presenting the results of processing by SmartMarker tools. A detailed description of the design of SmartMarkers will be given later.

25 Browser Extensions

Browser extensions for Microsoft's Internet Explorer 5 implement a standard API (Application Programming Interface) to enable developers to reserve space within the Browser's window for displaying links to additional functions and to access the Browser's content through its Document Object Model (DOM). In particular an area of the window is typically reserved as a so-called "Explorer Bar". The DOM is a programming interface specification which lets a programmer create and modify HTML pages and XML documents as program objects. The DOM describes the logical structure of documents and the way a document is accessed and interacted with. Individual HTML elements within a document are separately addressable objects having a separate representation within the DOM, and each object is accessible via its pointer. Navigating to a desired object involves navigating through a chain of object pointers.

40 Known uses of Browser extensions have included use as a mechanism for linking to other applications running on a remote server computer. They have also been used to enable an application to interact with Web page data content, but only in very limited ways such as highlighting a block of text or listing and displaying images. There has previously been

no generic mechanism on a Web client system providing the capability to process the semantic data content of all HTML-based Web pages.

SmartMarker-enabled Browser

5 Referring to Figure 1, the first embodiment of the invention comprises the following software components: a Web Browser 10; a SmartMarker Browser extension Dynamic Link Library (DLL) 30; a SmartMarker user interface DLL 40; a SmartMarker transformation layer 70; and a plurality of SmartMarker components 80 and their respective
10 SmartMarker algorithms 90. The SmartMarker Browser extension DLL and SmartMarker user interface DLL together provide access to the Web page data content 20 via the Browser's Document Object Model, enabling extraction of the data values of a set of data items selected by the user and enabling the extracted values to be sent to the SmartMarker.
15 transformation layer. The Smartmarker transformation layer interoperates with a SmartMarker component and associated SmartMarker algorithm to process the extracted data values. The results of the processing are returned to the SmartMarker user interface DLL 40 (via the transformation layer 70) which processes the results and applies them to the data content in the Browser. This will be described below in more detail.
20

25 The SmartMarker Browser extension DLL 30 implements the required ActiveX interfaces and services for an ActiveX Browser extension DLL, reserves user interface space within the Browser's window, acquires an ActiveX interface pointer to the Document Object Model for the Browser to provide access to the Browser's displayed data content, creates an instance of the SmartMarkers user interface ActiveX DLL 40, and displays the user interface defined by the user interface DLL in the reserved space.
30

35 The SmartMarker user interface DLL 40 is used in the creation of an instance of a SmartMarker Explorer Bar 50 and a SmartMarker palette 60 (see below), as well as its use in the accessing and extraction of data values from the Browser's data content and processing and applying the results of SmartMarker processing.

40 The SmartMarker Browser extension DLL 30 and SmartMarker user interface DLL 40 are implemented as self-registering ActiveX DLLs. Installation of the SmartMarkers Browser extension therefore involves the SmartMarker Browser extension DLL 30 and the SmartMarker user interface DLL 40 registering themselves in the Windows registry using standard methods. When the Browser is started it looks in the Windows registry for any Browser extension DLLs and lists those extensions on its Explorer Bar menu (accessible by selecting 'View' then 'Explorer Bar' from within the Browser). If the SmartMarker Browser extension is found, a SmartMarkers entry is added to the menu. This registry lookup and adding to the
45

Explorer Bar menu is standard implementation of the Microsoft Browser extension architecture.

5 A user can interact with the Browser using standard techniques to navigate to a Web page of interest. An input controller controls the position of a pointer on the display device in response to signals from the input devices, and the Browser responds to mouse related messages in order to identify and interpret selection of data items.

10 When a user selects SmartMarkers from the 'View' - 'Explorer Bar' menu, program code within the SmartMarker Browser extension DLL is invoked to:

15

- o reserve user interface space within the Browser's window for an Explorer Bar;
- o create an instance of the SmartMarkers Explorer Bar 50 using functions within the SmartMarker user interface DLL 40;
- o display the SmartMarkers Explorer Bar graphical user interface 50 in the reserved space;

20

- o access the DOM pointer for the Browser window;
- o pass the DOM pointer to the SmartMarker Explorer Bar 50 to provide access to the content 20 of the Browser window.

25 (Note that the Browser interface could alternatively be customised to enable the SmartMarker user interface to be displayed without using an explorer bar. Other options include tool bar buttons, a dedicated tool bar, menu options and pop-up windows.)

30 When displayed, the SmartMarker Explorer Bar 50 provides a first clickable button to enable the user to request display of a palette listing selectable SmartMarker operations. The SmartMarker Explorer Bar also provides a second button enabling the user to activate processing by a selected SmartMarker component of selected data content of a Web page displayed by the Browser. The SmartMarkers Explorer Bar will be displayed in subsequent Browser sessions until removed by the user by using the 'View' - 'Explorer Bar' menu.

40 When the user presses the 'Show SmartMarker palette' button in the SmartMarker Explorer Bar 50, program code within the SmartMarker user interface DLL 40 creates an instance 60 of a SmartMarker palette object and displays in a separate window a user interface provided by the SmartMarker palette object. The SmartMarker palette user interface lists available SmartMarker operations and listens for selection events, enabling the user to activate and deactivate a SmartMarker processing component, and then displays a key to help the user to interpret results

calculated by the activated SmartMarker and a button for requesting display of help for the activated SmartMarker.

When the user selects an individual SmartMarker operation from the list displayed in the SmartMarker palette 60, the respective SmartMarker object 80 is activated. If the user has not already done so, they then navigate to a Web page of interest. If the user then selects a block of data content in the Browser's window (using the standard Browser user interface for selection) and presses the 'SmartMark selected text' button in the SmartMarker Explorer Bar 50, the SmartMarker Explorer Bar 50 passes the DOM pointer for the Browser to the SmartMarker user interface DLL 40 and program code within the SmartMarker user interface DLL 40 is invoked to:

- o navigate from the DOM pointer for the Browser 10 to the DOM pointer for the selected content within the current document (i.e. stepping through a chain of pointers to addresses for particular objects from the Browser window object to the document object, and so on until the pointer to the selected data content is reached);
- o copy the HTML string corresponding to the selected data content into an appropriate area of memory (or alternatively, in a non-Java implementation, store the pointer);
- o parse the HTML string corresponding to the selected content to separate tokens such as words and numbers from HTML tags (implementation of parsing to strip values from strings is known in the art);
- o parse the separated tokens to strip trailing punctuation and other extraneous characters ("noise" such as currency symbols and thousands separators); and
- o send the tokens to the SmartMarker transformation layer.

When the SmartMarker transformation layer receives the 'cleaned' tokens, it formats and filters those tokens to meet the needs of the SmartMarker algorithm associated with the active SmartMarker. For example, some SmartMarker algorithms require only numeric data, while others require only string data. After formatting, the tokens are sent to the algorithm to be processed. This step is standard processing for SmartMarkers.

When the SmartMarker algorithm receives the tokens, it retrieves information about the current SmartMarker (obtaining its detailed properties) and then executes the method `Algorithm.calculateResults` to perform calculations on the data values of the tokens. The calculation is specific to the individual SmartMarker processing operation. For example, a SmartMarker algorithm might compare each token to a stored list of known keywords such as the names of partners or competitors, and

highlight matches. A second example is a spell checker. A third example is to compare each token to a database or dictionary of meanings of acronyms, or technical, legal, or business terms, and to annotate individual words when the mouse pointer overlays them. A fourth example is to access from a database, and to display in pop-up windows, information about products, companies or individual people identified within a Web page - the name of the product, company or person can be used as a key for a database lookup operation. A further example SmartMarker operation is to draw a business chart to aid analysis of numbers in a table or embedded in free-form text.

The results of the calculations are in two forms: emphasis results and annotation results. This performance of calculations to generate results which are relevant to the semantics of the data is standard SmartMarkers processing.

Emphasis results are used to highlight individual tokens. For example, a different emphasis might be used to highlight partners from that used to highlight competitors. Annotation results are used to provide additional information about a token. For example, in response to a mouse moving over the name of a competitor within a processed portion of text, a brief description of that competitor company might be displayed.

The results are returned to the SmartMarker user interface DLL through the SmartMarker transformation layer, and `endResults()` is called to notify the user interface DLL when a complete set of results has been provided. On receipt, these results are processed as follows:

- For each token that has any results, an HTML `...` tag pair is constructed to wrap that token. The `...` tag set is a standard HTML tag set used to separate one section of a Web page from another. It enables application of a set of attributes such as color or font to a range of text between the tag pairs.
- Constructing these tags for individual tokens enables the setting of specific attributes representing the results of the SmartMarker processing for individual number values or text words. In the known prior art, it is not typical for individual words or number values within a page of HTML to have an HTML `...` tag pair associated with them.
- For emphasis results, a style parameter is added to the `` tag in the form `STYLE='background-color:#rrggbb'` where rr, gg, bb represent the red, green, and blue components of a color calculated for this token by the SmartMarker transformation layer from the emphasis result calculated by the SmartMarker algorithm. The style

parameter is a standard CSS (Cascading style sheet) attribute recognized by the web browser as a request to change the presentation of a block of content.

5 o For annotation results, a segment of dynamic HTML is added to the tag in the form onmouseover=' window.status="xxx"' onmouseout='window.status=""' where xxx represents the annotation string calculated for this token by the SmartMarker algorithm. The dynamic HTML is a standard tag attribute recognized by the Web

10 o Browser as a request to display the specified string in the browser status area when the user moves the mouse over this token and to clear the status area when the mouse is moved elsewhere. The Browser status area is a fixed area of the Browser window reserved for status messages.

15 o As noted previously, annotation results may equally be presented as pop-up boxes ('Hover Help' style) by adding TITLE='xxx' to the SPAN tag pair, instead of status line messages.

20 When the results have been processed, the SmartMarker user interface DLL applies the results to the content in the Browser as follows:

25 o The tokens wrapped in .. tag pairs are merged with the copy of the original HTML of the selected data content which is held in memory, by replacing the original unwrapped tokens to create a replacement HTML string.

30 (In an alternative non-Java implementation, if no copy of the extracted HTML was stored in memory and only a pointer was stored, a set of replacement HTML strings comprising the wrapped tokens is associated with the stored pointer).

35 o A standard document object model service (pasteHTML) is then called to paste the replacement HTML string into the currently selected Web page.

40 o The Browser immediately reinterprets the replaced HTML and honours the tag parameters by applying background color and mouse-over support to the tokens modified in response to emphasis and annotation results calculated by the algorithm.

45 Thus, from a user's perspective, selection of a specific SmartMarker operation from a list displayed in a palette, selection of a set of data items such as a block of text or a set of numerical values from a Web page, and pressing the 'SmartMark selected text' button simply

results in the selected area of the Web page being updated with highlights and annotations.

Figure 2 shows an example display screen including a Browser window 100, SmartMarker explorer bar 50 including selectable buttons 110, and a SmartMarkers palette 60 displayed in a separate window. The user has selected the 'Show SmartMarker palette' button within the explorer bar, then selected the 'News Reader' SmartMarker operation from the palette, navigated to a Web page of interest and selected text of interest. On selection of the 'SmartMark selected text' button, the portion of selected text is extracted, text words are separated from HTML tags and are sent to the respective SmartMarker. The SmartMarker compares each word in the set with a list of company names to identify company names, matches the company names to particular categories of company, and generates an annotation string and emphasis colour for each matched company name. These annotation and emphasis results are then applied to the HTML content and displayed by the Browser.

A particular example of the stages of processing of a block of text selected from a Browser's data content will now be given. Assume that a user of a Web client data processing apparatus, which includes software components implementing the present invention, selects a SmartMarker operation which identifies acronyms and generates an emphasis colour result and an annotation string for each recognised acronym. The annotation string is an expansion of the acronym. Assume that the user selects the following text content:

3 important memory types are RAM, DRAM, and ROM.

This data content in tokenized form is:

```
3
"
important
"
memory
"
types
"
are
"
RAM,
"
"
DRAM,
"
"
and
"
"
```

ROM.

The data content when filtered for a SmartMarker that requires text string tokens is:

5
 important
 memory
 types
 are
 10 RAM,
 DRAM,
 and
 ROM.

15 The data tokens when stripped of noise are:

important
 memory
 types
 are
 20 RAM
 DRAM
 and
 ROM

25 Some tokens are then modified by SmartMarker results as follows:

30 <SPAN style='background-color:ff0000'
 onmouseover='window.status="Random access memory"'
 onmouseout='window.status=""'>RAM,
 <SPAN style='background-color:ff0000'
 onmouseover='window.status="Dynamic random access memory"'
 onmouseout='window.status=""'>DRAM,
 <SPAN style='background-color:ff0000'
 onmouseover='window.status="Read only memory"'
 onmouseout='window.status=""'>ROM.

When reassembled and pasted back into the Browser, the modified HTML is:

40 3 important memory types are <SPAN style='background-color:ff0000'
 onmouseover='window.status="Random access memory"'
 onmouseout='window.status=""'>RAM, <SPAN style='background-
 color:ff0000' onmouseover='window.status="Dynamic random access
 memory"' onmouseout='window.status=""'>DRAM, and <SPAN
 style='background-color:ff0000' onmouseover='window.status="Read
 only memory"' onmouseout='window.status=""'>ROM.

SmartMarkers Design

In the preferred embodiment of the present invention, SmartMarker components are implemented as "invisible" Java BeansTM. Java Beans are self-contained software component classes written in the JavaTM programming language and which can be combined with other Java Beans to perform a task when run on a computer with the support of a Java Virtual Machine. Java components are known in the art to execute as a set of objects within a Java Virtual Machine. Objects such as SmartMarker components 80,90 within the same JVM can establish communication connections with each other by calling standard Java static methods on the classes associated with other objects. Static methods are methods associated with a class rather than an instance of the class. "Invisible" in the above context means that the individual objects are represented at run time by a menu list within a palette displayed on a display device rather than by separate screen icons.

This palette 60 is itself implemented as a "visible" Java Bean which operates within the same JVM as the SmartMarker Java Beans 80,90. The SmartMarker operation selection palette Java Bean thus allows the user at run time to select a SmartMarker operation from a list using a mouse. The SmartMarker palette object identifies selections in response to mouse events. The use of a palette ensures that the run time representation of the SmartMarker operations takes up minimal screen space and indicates to end users that the SmartMarker operations are conceptually associated with each other and are distinct from any unrelated operations. This mechanism for item selection within a GUI and various alternatives are well known in the art.

Each SmartMarker is responsible for a specific data interpretation operation or data generation operation that processes a set of data values supplied by a SmartMarker-enabled Browser and calculates a set of results. The results may be, for example, emphasis colours and/or annotation, as will be described below. For example, an 'Average' SmartMarker determines whether numerical data values within a selected set are above, close to or below an average and then either selects emphasis colours to apply to the values to represent these categories and/or generates annotation strings interpreting individual values; whereas the 'Trend' SmartMarker calculates whether values are higher or lower than the previous value and then selects colours and/or generates an annotation result to reflect this calculation result.

The logic that defines a SmartMarker operation is coded in a SmartMarker algorithm object 90 associated with the SmartMarker Java Bean 80. A SmartMarker Java Bean represents a single SmartMarker within the palette Java Bean. Each such bean has a number of properties including title and algorithm name. This algorithm name property defines the class

name of the algorithm object 90 to be associated with the SmartMarker Java Bean. During initialisation of the SmartMarker Java Bean 80, this class name is used to construct an instance of an algorithm object 90 to be executed later when a user selects the SmartMarker Java Bean 80 and selects a set of values.

Algorithm objects 90 implement SmartMarker interface 'Algorithm' which interface defines a set of methods that will be called on an algorithm by functions within the SmartMarker transformation layer 70 after transforming the data supplied by the user interface DLL 40.

The implementation of SmartMarker algorithms 90 as separate objects from their associated SmartMarker objects 80 enables certain SmartMarkers with different properties to use the same SmartMarker algorithm object 90 for the generation of different interpretation results. Nevertheless, many SmartMarker algorithm objects 90 are only associated with a single SmartMarker object 80 (and so in alternative embodiments of the invention the SmartMarker and SmartMarker algorithm for these operations could be implemented as a single object).

To implement the Algorithm interface, an algorithm object:

- o implements method `getSmartMarkerAttributes()` to enable other parts of the SmartMarker operation model to retrieve information about the current SmartMarker (its detailed properties or attributes such as a set of direction-of-processing options displayed in the user interface which may be relevant for tabulated data values);
- o implements method `getRequiredDataTypes()` to indicate whether it requires string or numeric data, or some other data type such as links or other HTML constructs;
- o implements method `calculateResults()` to:
 - accept a value set containing a set of selected values supplied by the SmartMarker user interface DLL via the SmartMarkers transformation layer;
 - extract value objects of a specific data partition from that value set (e.g. the partition may be a row, column, or block within a table of application data values, this partition being determined by a processing Direction option within the properties of the current SmartMarker);
 - perform calculations upon these extracted values to generate a set of results;
- o optionally, an algorithm object then calls methods on some or all of the value objects to set an associated emphasis colour result

5 and/or calls methods on some or all of the value objects to set an associated annotation result. These *setEmphasis* and *setAnnotation* methods are implemented within a class called *ValueBase*, which is the base class for classes that represent values or sets of values passed to the algorithm (*Value*, *ValuePartition*, and *ValueSet*), and allow the algorithm to set a result on a specific value or set of values.

10 The above mentioned ability for the user to select different directions of processing is useful when data is organised in a structured form such as a table.

15 When the algorithm has completed execution of method *calculateResults*, the results are returned to the SmartMarker user interface DLL as described above.

20 The following methods are implemented by the SmartMarker user interface DLL (these methods are defined in mandatory application interface *SmartMarkerEnabler*, and optional application interfaces *EmphasisColor* and *Annotation*):

SmartMarkerEnabler.startResults

25 Notifies the user interface DLL that a set of results are about to be supplied, prompting the user interface DLL to start building a replacement HTML string.

EmphasisColor.applyEmphasisColor

30 Called once for each value supplied to the algorithm upon which the algorithm has called *ValueBase.setEmphasis*. Implemented by the user interface DLL to create a CSS style parameter to be added to the tag generated for the associated value in the data content in the replacement HTML.

35 *Annotation.applyAnnotation*

40 Called once for each value supplied to the algorithm upon which the algorithm has called *ValueBase.setAnnotation*. Implemented by the user interface DLL to create a DHTML string to be added to the tag generated for the associated value in the data content in the replacement HTML. In an implementation using a 'Hover Help' style, each such annotation string may subsequently be displayed in response to subsequent mouse proximity to the associated value.

45 *SmartMarkerEnabler.endResults*

Notifies the user interface DLL that all results have been sent, prompting the user interface DLL to complete generation of the modified HTML and to paste that HTML into the Browser's window to replace the original selected HTML.

5

As well as calling methods on a SmartMarker algorithm object, and implementing methods for setting annotations and emphasis colours, the respective SmartMarker object 80 has a set of properties including its name and a number of options including:

10

- o the type of partition of a selected data set to be analysed (e.g. single value, row, column, block),
- o a set of emphasis colours to be applied, for example red, yellow and green,
- o a set of text strings used as a key to interpret emphasis colours (i.e. each text string being applicable to one category of possible results).

20

Thus, a SmartMarker algorithm 90 operates upon values supplied by the Browser extension components 30,40,50 via the SmartMarker transformation layer 70 and returns a set of calculated results associated with those values which are used by the SmartMarker algorithm object 90 to generate interpretation annotation strings and emphasis colours which are returned to the Browser extension components.

25

Within the algorithm, each supplied value is represented by an instance of class Value. Class Value contains both the actual value and keys (such as row and column numbers) that enable the Browser extension components to recognize the source of the value. For algorithms which process numerical data values displayed in tabulated form in the Browser's window, values are grouped into partitions where each partition represents a row, column, or block. Each partition is represented as an instance of class ValuePartition. Value partitions are grouped into a single value set represented by an instance of class ValueSet.

30

An algorithm may return several types of results including colour changes, and annotations. To return a result for a value, the algorithm calls an appropriate method on an instance of class Value. To return a result for all values in a partition, the algorithm calls an appropriate method on an instance of class ValuePartition. To return a result for all values, the algorithm calls an appropriate method on an instance of class ValueSet.

40

45

The respective methods for generating annotation strings are:

```
5      Value.setAnnotation(String annotation)
      ValuePartition.setAnnotation(String annotation)
      ValueSet.setAnnotation(String annotation)
```

For typical operations, a SmartMarker algorithm object 90 will return the results for all values in a selected set, which the Browser extension components process and apply to the Browser content as described above. Display of annotation strings for individual data values then requires subsequent data value selection by movement of the mouse pointer to overlay the relevant data item within the Web page (which may be a cell of a table of numbers or a word within a portion of text). The correlation between mouse pointer position and individual words is handled by the Browser - for example, the Browser is instructed, by the DHTML onmouseover=' window.status="xxx"' which was added to the ... group, to display a status message whenever the mouse is over the word in a tag group.

20 The setting of annotation strings comprises mapping each numerical result of the calculation performed by the SmartMarker algorithm to one of a set of defined result categories specific to the current SmartMarker operation. Each category has an associated text string held as a property of the SmartMarker object. For numeric SmartMarker operations, the 25 setAnnotation method typically includes generating a text string which combines the text predefined for the identified category with the numerical result of the calculation. Examples of this generation will be described below under 'Example Operations'.

30 Example Operations

As noted previously, a great many different processing operations can be implemented as SmartMarkers, including interpretation operations such as calculating for each value of a selected set whether the value is 35 above or below the average for the set and then marking the values accordingly, or identifying and displaying a ranking of a set of data values.

40 US 5,896,491 (which is incorporated herein by reference) described interpretation operations in which a colour code or other emphasis is applied to individual data cells of a selected set of data cells within an application program. The colour or other applied emphasis represents the relation of each value to the other values in the set. For example, a Rank operation determines an ordering of data values by their magnitude, and the results of the operation may be presented to the user by using 45 different shades of blue for the background of the cell containing each

data value in the selected set, with the largest value having the deepest blue background. US 5,896,491 mentions that many forms of emphasis are possible as alternatives to background colours.

5 To implement the Rank operation, the processing program compares all of the values in the set with each other, or compares each with a common value, to determine a ranking by magnitude, and then selects an appropriate shading or other emphasis to apply to each value to represent the relative values. US 5,896,491 disclosed implementing this by
10 assigning values to percentile bands with all values in a given band being given the respective shading. The calculation thus preferably determines an actual ranking for each value but this is then represented to the user in a simplified, easily interpreted form.

15 In some circumstances, it will be highly desirable for the user to be able to confirm the actual position in the ranking of an individual value in addition to (or instead of) being presented with the very easily interpretable colour coding. For example, a user may wish for confirmation that a selected value is within the top 10% of the set of
20 selected values, or that it is ranked third out of 50 values.

25 The present invention makes possible the provision of such a confirmation to the user of a Web Browser, using the mechanism for implementation of interpretation and data generation operations described in UK patent application GB 9816098.9. This mechanism is capable of presenting per-value data interpretation information to a computer user. That is, as an alternative or addition to sending to the application component an identification of the required emphasis to be applied by the
30 Browser, a SmartMarker processing component according to the invention includes methods for providing annotation strings to the application component, the annotation strings describing the relation of a specific selected value to other values within a particular partition of a selected data set. The application component then displays an annotation string in response to user selection of a respective data value, such as
35 selection by movement of the mouse pointer over the cell containing the value.

40 In association with the Rank operation described above, the annotation strings which can be returned to the application component for display alongside individual selected values may be strings which simply put into words the same information as is represented by colour codes, such as "within top 10% of values in set", "within bottom 10% of values in set", etc. Alternatively, the strings may be more detailed and give numerical information which cannot be represented by applying colour or
45 other emphasis, such as: "ranked 3rd in row, 25 below highest".

5 As another example, an Average SmartMarker operation determines an arithmetic mean for a selected partition of a selected set of data and then compares individual values to the mean. The numerical results of this comparison are assigned to one of a number of categories. These categories may be, for example, the three categories "greater than mean", "within 10% of mean", and "less than mean". Each category has corresponding text stored in association therewith for inclusion within annotation text strings.

10 If a row of a selected data set has a mean of "M", and an individual value "N" within the selected set is determined to be greater than the computed mean for a particular partition ("Partition1", which is a row, column or block of data values - as specified within the properties of the SmartMarker by selecting the "Direction" tab of the 15 palette) by an amount "deltaString", then the Average SmartMarker algorithm initially returns the two intermediate results M and deltaString, and in this case the intermediate results for value N are mapped to the "greater than" category. An annotation string is then generated by the setAnnotation method. This uses the intermediate results M and deltaString and stored text associated with the "greater than" 20 result category as follows:

```
annotation = "N is deltaString greater than M (Partition1 mean)".
```

25 This annotation string is then returned to the SmartMarker user interface DLL for applying to the Browser's data content from which the data values were taken. An annotation text string is preferably returned for each value within the selected set of data values and the strings are stored by the Browser.

30 The Browser displays an annotation string in response to the mouse pointer moving over the respective data value. The Average SmartMarker has also returned to the Browser a set of annotation strings and a series 35 of selected emphasis colours representing the different categories referred to above. The Browser has displayed on the screen the emphasis colours for the block of selected data values. The emphasis colours are obtained using the same results of the comparison between individual data values and the other values in the row, and the same categorisation. However, instead of generating annotation strings using text associated 40 with different result categories, the emphasis operation obtains a colour associated with a result category and returns this to the application component for display.

45 As the mouse pointer is moved to a different data value within the same partition of the selected data set, a different one of the

annotation strings (included in the replaced HTML following generation of the strings by the Average SmartMarker) is displayed.

Thus, the methods applying emphasis colours for the Average operation provide a visually intuitive categorisation of specified partitions of a selected set of results and the methods for presenting annotation information provide supplementary information which assists user interpretation and saves the user from mental arithmetic.

An annotation operation can thus be provided in association with an emphasis-applying operation implemented by a single processing component, for processing data values supplied by a Web Browser and returning the generated results to the Browser. Alternatively, annotation may be implemented independently of emphasis.

A third example is a Compare operation, which compares each value with the first value in the row, column, or block (which partition being determined by the 'Direction' property of the SmartMarker). The positioning of the mouse pointer over a value is identified by the Browser which then presents the respective stored annotation string. The annotation strings generated by the SmartMarker object and stored by the Browser in response to selecting this Compare SmartMarker operation and selecting a set of data are of the form:

annotation = "N is deltaString greater than N1 (first in Partition1)"

where deltaString is in this case the result of subtracting the first value in the row (N1) from the current selected value (N). Partition1 is, once again, the row of the selected data set which contains the current selected value. Emphasis colours represent whether each value is greater than or less than the first value in the row.

A further example SmartMarker operation is a Trend operation, which compares each value in a selected set with the previous value in the partition (in this case, a row, ordered from left to right). Emphasis colours indicate whether each value in a row of a selected set of values is less than or greater than the previous value, and the annotation result provides alphanumeric interpretation information describing the difference between the two values. This can be very useful in aiding the interpretation of trends within data.

A further example SmartMarker operation compares each value in a selected set with a specific reference value and highlights only those results which are greater than the reference value. In this operation, data values have been selected as a block and the numerical intermediate

result of the calculation is a number by which each individual value exceeds the reference value. The SmartMarker assigns positive results to a "greater than" category and generates an annotation string such as is shown in the example of the Figure. For this example, no emphasis colour is applied to data values which had a negative result when compared with the reference value, simplifying the presentation to the user.

5 Alternatively, it is possible to mark all values to remind the user which set of values was selected and to apply different emphasis colours to positive and negative results.

10 Further example operations include a Total SmartMarker which uses annotation to display the total of all values in each row or column. This is an example in which annotation can provide useful interpretation information which is relevant to the current operation and the current 15 data values of a selected set and which is implemented independently of emphasis colours.

A WordCount SmartMarker uses annotation to display a count of the words in a selected block of text.

20 Statistical analysis SmartMarker and chart generator SmartMarker are examples of data generator SmartMarker operations which may also benefit from interpretative annotation.

25 It will be clear to persons skilled in the art that many different operations can be implemented using methods and mechanisms according to the present invention. In some cases the display of interpretation information will be most helpful to users if implemented together with applying per-value emphasis to represent categories of results, but in other cases the display of per-value interpretation information will be most useful if implemented as flyover text or some other media without 30 additional emphasis. The per-value interpretation information can take many forms and even be presented via different media types as alternatives to the screen display of text strings described in detail above.

35 In view of algorithms used by SmartMarkers being implemented as software components which are executable independently of the Web Browser and are independent of individual Web pages, SmartMarker objects can be enhanced and additional SmartMarkers can be provided without having to 40 modify the Web Browser or individual Web pages. The invention is significant in its provision of access to interpretation and analysis facilities from a Web Browser to enable semantic data processing on a Web client system of Web page content.

5 While the present invention has been described in detail with reference to Internet Explorer 5 and Browser extensions, it will be appreciated by persons skilled in the art that the invention is equally applicable to other Web Browsers implementing a different user interface design. The above descriptions of particular designs of user interface for a Browser extension implementation using an Explorer Bar is exemplary only, and it will be recognised that many alternative interface designs are possible within the scope of the invention.

10 The specific implementation described in detail presents the results of processing of Web page content as modified HTML displayed within the Browser's Web page display area, or displays the results elsewhere in the Browser's window. Alternatively, the processing components may present their results via some other mechanism (e.g. 15 generating an audio result which is then played back via a speaker) while the original Web page is being displayed. In this case, processing results are still preferably provided to the Browser by the processing components and then the Browser or an associated presentation handler component manages presentation of different data types (text, images, 20 audio) via appropriate presentation devices (display screen, speakers). In either case, the provision of results to the Browser does not involve downloading of a replacement Web page from a server. Such non-visual presentation of processing results is particularly relevant if the device for accessing Web page content uses non-visual interfaces for data 25 access.

CLAIMS

5 1. A set of components for cooperating with a Web Browser to process the semantic data content of Web pages, wherein the Web Browser implements functions to present Web page contents to a user of a Web client data processing apparatus, and wherein the set of components are operable to run on the Web client data processing apparatus together with the Web Browser, the set of components including:

10 a data content transfer component, responsive to user selection of a desired semantic processing operation and user selection of a set of data items within a presented Web page, for extracting the set of data items from the Web page; and

15 a set of data content processing components, each for processing individual data items of a set of data items extracted from any one of multiple Web pages, for performing at the Web client data processing apparatus the selected processing operation on the individual data items of the selected set, to generate a result which is dependent on the semantics of the selected set of data items;

20 wherein the data content transfer component includes means for providing the generated result to the Web Browser for presentation to the user.

25 2. A set of components according to claim 1, wherein the data content transfer component implements functions for:

30 accessing the user-selected portion of the markup language content of a Web page, separating individual data items from markup tags within the selected portion, and passing the separated data items to one of said set of data content processing components; and,

35 in response to receipt of a generated result from said data content processing component, generating a modified version of the portion of the markup language Web page content, and providing said modified portion to the Web Browser to replace the extracted portion of the presented Web page.

40 3. A set of components according to claim 2, wherein the separating of data items comprises parsing the selected portion of the markup language content to separate data items from markup tags, and parsing the separated data items to remove any extraneous characters.

4. A set of components according to claim 2 or claim 3, wherein the generating of a modified version of the portion of the markup language Web page content comprises constructing a markup language separator tag pair for each data item for which results are returned, separately wrapping each of said data items using said tag pair and applying the results to the wrapped data items.

5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100 105 110 115 120 125 130 135 140 145 150 155 160 165 170 175 180 185 190 195 200 205 210 215 220 225 230 235 240 245 250 255 260 265 270 275 280 285 290 295 300 305 310 315 320 325 330 335 340 345 350 355 360 365 370 375 380 385 390 395 400 405 410 415 420 425 430 435 440 445 450 455 460 465 470 475 480 485 490 495 500 505 510 515 520 525 530 535 540 545 550 555 560 565 570 575 580 585 590 595 600 605 610 615 620 625 630 635 640 645 650 655 660 665 670 675 680 685 690 695 700 705 710 715 720 725 730 735 740 745 750 755 760 765 770 775 780 785 790 795 800 805 810 815 820 825 830 835 840 845 850 855 860 865 870 875 880 885 890 895 900 905 910 915 920 925 930 935 940 945 950 955 960 965 970 975 980 985 990 995 1000 1005 1010 1015 1020 1025 1030 1035 1040 1045 1050 1055 1060 1065 1070 1075 1080 1085 1090 1095 1100 1105 1110 1115 1120 1125 1130 1135 1140 1145 1150 1155 1160 1165 1170 1175 1180 1185 1190 1195 1200 1205 1210 1215 1220 1225 1230 1235 1240 1245 1250 1255 1260 1265 1270 1275 1280 1285 1290 1295 1300 1305 1310 1315 1320 1325 1330 1335 1340 1345 1350 1355 1360 1365 1370 1375 1380 1385 1390 1395 1400 1405 1410 1415 1420 1425 1430 1435 1440 1445 1450 1455 1460 1465 1470 1475 1480 1485 1490 1495 1500 1505 1510 1515 1520 1525 1530 1535 1540 1545 1550 1555 1560 1565 1570 1575 1580 1585 1590 1595 1600 1605 1610 1615 1620 1625 1630 1635 1640 1645 1650 1655 1660 1665 1670 1675 1680 1685 1690 1695 1700 1705 1710 1715 1720 1725 1730 1735 1740 1745 1750 1755 1760 1765 1770 1775 1780 1785 1790 1795 1800 1805 1810 1815 1820 1825 1830 1835 1840 1845 1850 1855 1860 1865 1870 1875 1880 1885 1890 1895 1900 1905 1910 1915 1920 1925 1930 1935 1940 1945 1950 1955 1960 1965 1970 1975 1980 1985 1990 1995 2000 2005 2010 2015 2020 2025 2030 2035 2040 2045 2050 2055 2060 2065 2070 2075 2080 2085 2090 2095 2100 2105 2110 2115 2120 2125 2130 2135 2140 2145 2150 2155 2160 2165 2170 2175 2180 2185 2190 2195 2200 2205 2210 2215 2220 2225 2230 2235 2240 2245 2250 2255 2260 2265 2270 2275 2280 2285 2290 2295 2300 2305 2310 2315 2320 2325 2330 2335 2340 2345 2350 2355 2360 2365 2370 2375 2380 2385 2390 2395 2400 2405 2410 2415 2420 2425 2430 2435 2440 2445 2450 2455 2460 2465 2470 2475 2480 2485 2490 2495 2500 2505 2510 2515 2520 2525 2530 2535 2540 2545 2550 2555 2560 2565 2570 2575 2580 2585 2590 2595 2600 2605 2610 2615 2620 2625 2630 2635 2640 2645 2650 2655 2660 2665 2670 2675 2680 2685 2690 2695 2700 2705 2710 2715 2720 2725 2730 2735 2740 2745 2750 2755 2760 2765 2770 2775 2780 2785 2790 2795 2800 2805 2810 2815 2820 2825 2830 2835 2840 2845 2850 2855 2860 2865 2870 2875 2880 2885 2890 2895 2900 2905 2910 2915 2920 2925 2930 2935 2940 2945 2950 2955 2960 2965 2970 2975 2980 2985 2990 2995 3000 3005 3010 3015 3020 3025 3030 3035 3040 3045 3050 3055 3060 3065 3070 3075 3080 3085 3090 3095 3100 3105 3110 3115 3120 3125 3130 3135 3140 3145 3150 3155 3160 3165 3170 3175 3180 3185 3190 3195 3200 3205 3210 3215 3220 3225 3230 3235 3240 3245 3250 3255 3260 3265 3270 3275 3280 3285 3290 3295 3300 3305 3310 3315 3320 3325 3330 3335 3340 3345 3350 3355 3360 3365 3370 3375 3380 3385 3390 3395 3400 3405 3410 3415 3420 3425 3430 3435 3440 3445 3450 3455 3460 3465 3470 3475 3480 3485 3490 3495 3500 3505 3510 3515 3520 3525 3530 3535 3540 3545 3550 3555 3560 3565 3570 3575 3580 3585 3590 3595 3600 3605 3610 3615 3620 3625 3630 3635 3640 3645 3650 3655 3660 3665 3670 3675 3680 3685 3690 3695 3700 3705 3710 3715 3720 3725 3730 3735 3740 3745 3750 3755 3760 3765 3770 3775 3780 3785 3790 3795 3800 3805 3810 3815 3820 3825 3830 3835 3840 3845 3850 3855 3860 3865 3870 3875 3880 3885 3890 3895 3900 3905 3910 3915 3920 3925 3930 3935 3940 3945 3950 3955 3960 3965 3970 3975 3980 3985 3990 3995 4000 4005 4010 4015 4020 4025 4030 4035 4040 4045 4050 4055 4060 4065 4070 4075 4080 4085 4090 4095 4100 4105 4110 4115 4120 4125 4130 4135 4140 4145 4150 4155 4160 4165 4170 4175 4180 4185 4190 4195 4200 4205 4210 4215 4220 4225 4230 4235 4240 4245 4250 4255 4260 4265 4270 4275 4280 4285 4290 4295 4300 4305 4310 4315 4320 4325 4330 4335 4340 4345 4350 4355 4360 4365 4370 4375 4380 4385 4390 4395 4400 4405 4410 4415 4420 4425 4430 4435 4440 4445 4450 4455 4460 4465 4470 4475 4480 4485 4490 4495 4500 4505 4510 4515 4520 4525 4530 4535 4540 4545 4550 4555 4560 4565 4570 4575 4580 4585 4590 4595 4600 4605 4610 4615 4620 4625 4630 4635 4640 4645 4650 4655 4660 4665 4670 4675 4680 4685 4690 4695 4700 4705 4710 4715 4720 4725 4730 4735 4740 4745 4750 4755 4760 4765 4770 4775 4780 4785 4790 4795 4800 4805 4810 4815 4820 4825 4830 4835 4840 4845 4850 4855 4860 4865 4870 4875 4880 4885 4890 4895 4900 4905 4910 4915 4920 4925 4930 4935 4940 4945 4950 4955 4960 4965 4970 4975 4980 4985 4990 4995 5000 5005 5010 5015 5020 5025 5030 5035 5040 5045 5050 5055 5060 5065 5070 5075 5080 5085 5090 5095 5100 5105 5110 5115 5120 5125 5130 5135 5140 5145 5150 5155 5160 5165 5170 5175 5180 5185 5190 5195 5200 5205 5210 5215 5220 5225 5230 5235 5240 5245 5250 5255 5260 5265 5270 5275 5280 5285 5290 5295 5300 5305 5310 5315 5320 5325 5330 5335 5340 5345 5350 5355 5360 5365 5370 5375 5380 5385 5390 5395 5400 5405 5410 5415 5420 5425 5430 5435 5440 5445 5450 5455 5460 5465 5470 5475 5480 5485 5490 5495 5500 5505 5510 5515 5520 5525 5530 5535 5540 5545 5550 5555 5560 5565 5570 5575 5580 5585 5590 5595 5600 5605 5610 5615 5620 5625 5630 5635 5640 5645 5650 5655 5660 5665 5670 5675 5680 5685 5690 5695 5700 5705 5710 5715 5720 5725 5730 5735 5740 5745 5750 5755 5760 5765 5770 5775 5780 5785 5790 5795 5800 5805 5810 5815 5820 5825 5830 5835 5840 5845 5850 5855 5860 5865 5870 5875 5880 5885 5890 5895 5900 5905 5910 5915 5920 5925 5930 5935 5940 5945 5950 5955 5960 5965 5970 5975 5980 5985 5990 5995 6000 6005 6010 6015 6020 6025 6030 6035 6040 6045 6050 6055 6060 6065 6070 6075 6080 6085 6090 6095 6100 6105 6110 6115 6120 6125 6130 6135 6140 6145 6150 6155 6160 6165 6170 6175 6180 6185 6190 6195 6200 6205 6210 6215 6220 6225 6230 6235 6240 6245 6250 6255 6260 6265 6270 6275 6280 6285 6290 6295 6300 6305 6310 6315 6320 6325 6330 6335 6340 6345 6350 6355 6360 6365 6370 6375 6380 6385 6390 6395 6400 6405 6410 6415 6420 6425 6430 6435 6440 6445 6450 6455 6460 6465 6470 6475 6480 6485 6490 6495 6500 6505 6510 6515 6520 6525 6530 6535 6540 6545 6550 6555 6560 6565 6570 6575 6580 6585 6590 6595 6600 6605 6610 6615 6620 6625 6630 6635 6640 6645 6650 6655 6660 6665 6670 6675 6680 6685 6690 6695 6700 6705 6710 6715 6720 6725 6730 6735 6740 6745 6750 6755 6760 6765 6770 6775 6780 6785 6790 6795 6800 6805 6810 6815 6820 6825 6830 6835 6840 6845 6850 6855 6860 6865 6870 6875 6880 6885 6890 6895 6900 6905 6910 6915 6920 6925 6930 6935 6940 6945 6950 6955 6960 6965 6970 6975 6980 6985 6990 6995 7000 7005 7010 7015 7020 7025 7030 7035 7040 7045 7050 7055 7060 7065 7070 7075 7080 7085 7090 7095 7100 7105 7110 7115 7120 7125 7130 7135 7140 7145 7150 7155 7160 7165 7170 7175 7180 7185 7190 7195 7200 7205 7210 7215 7220 7225 7230 7235 7240 7245 7250 7255 7260 7265 7270 7275 7280 7285 7290 7295 7300 7305 7310 7315 7320 7325 7330 7335 7340 7345 7350 7355 7360 7365 7370 7375 7380 7385 7390 7395 7400 7405 7410 7415 7420 7425 7430 7435 7440 7445 7450 7455 7460 7465 7470 7475 7480 7485 7490 7495 7500 7505 7510 7515 7520 7525 7530 7535 7540 7545 7550 7555 7560 7565 7570 7575 7580 7585 7590 7595 7600 7605 7610 7615 7620 7625 7630 7635 7640 7645 7650 7655 7660 7665 7670 7675 7680 7685 7690 7695 7700 7705 7710 7715 7720 7725 7730 7735 7740 7745 7750 7755 7760 7765 7770 7775 7780 7785 7790 7795 7800 7805 7810 7815 7820 7825 7830 7835 7840 7845 7850 7855 7860 7865 7870 7875 7880 7885 7890 7895 7900 7905 7910 7915 7920 7925 7930 7935 7940 7945 7950 7955 7960 7965 7970 7975 7980 7985 7990 7995 8000 8005 8010 8015 8020 8025 8030 8035 8040 8045 8050 8055 8060 8065 8070 8075 8080 8085 8090 8095 8100 8105 8110 8115 8120 8125 8130 8135 8140 8145 8150 8155 8160 8165 8170 8175 8180 8185 8190 8195 8200 8205 8210 8215 8220 8225 8230 8235 8240 8245 8250 8255 8260 8265 8270 8275 8280 8285 8290 8295 8300 8305 8310 8315 8320 8325 8330 8335 8340 8345 8350 8355 8360 8365 8370 8375 8380 8385 8390 8395 8400 8405 8410 8415 8420 8425 8430 8435 8440 8445 8450 8455 8460 8465 8470 8475 8480 8485 8490 8495 8500 8505 8510 8515 8520 8525 8530 8535 8540 8545 8550 8555 8560 8565 8570 8575 8580 8585 8590 8595 8600 8605 8610 8615 8620 8625 8630 8635 8640 8645 8650 8655 8660 8665 8670 8675 8680 8685 8690 8695 8700 8705 8710 8715 8720 8725 8730 8735 8740 8745 8750 8755 8760 8765 8770 8775 8780 8785 8790 8795 8800 8805 8810 8815 8820 8825 8830 8835 8840 8845 8850 8855 8860 8865 8870 8875 8880 8885 8890 8895 8900 8905 8910 8915 8920 8925 8930 8935 8940 8945 8950 8955 8960 8965 8970 8975 8980 8985 8990 8995 9000 9005 9010 9015 9020 9025 9030 9035 9040 9045 9050 9055 9060 9065 9070 9075 9080 9085 9090 9095 9100 9105 9110 9115 9120 9125 9130 9135 9140 9145 9150 9155 9160 9165 9170 9175 9180 9185 9190 9195 9200 9205 9210 9215 9220 9225 9230 9235 9240 9245 9250 9255 9260 9265 9270 9275 9280 9285 9290 9295 9300 9305 9310 9315 9320 9325 9330 9335 9340 9345 9350 9355 9360 9365 9370 9375 9380 9385 9390 9395 9400 9405 9410 9415 9420 9425 9430 9435 9440 9445 9450 9455 9460 9465 9470 9475 9480 9485 9490 9495 9500 9505 9510 9515 9520 9525 9530 9535 9540 9545 9550 9555 9560 9565 9570 9575 9580 9585 9590 9595 9600 9605 9610 9615 9620 9625 9630 9635 9640 9645 9650 9655 9660 9665 9670 9675 9680 9685 9690 9695 9700 9705 9710 9715 9720 9725 9730 9735 9740 9745 9750 9755 9760 9765 9770 9775 9780 9785 9790 9795 9800 9805 9810 9815 9820 9825 9830 9835 9840 9845 9850 9855 9860 9865 9870 9875 9880 9885 9890 9895 9900 9905 9910 9915 9920 9925 9930 9935 9940 9945 9950 9955 9960 9965 9970 9975 9980 9985 9990 9995 9999

processing component which is capable of processing data items of a set of data items extracted from any one of multiple Web pages;

5 providing the generated result to the Web Browser for presentation to the user.

10 11. A computer program for running on a Web client data processing apparatus and cooperating with a Web Browser to process the semantic data content of Web pages, wherein the Web Browser implements functions to present Web page contents to a user of the Web client data processing apparatus and wherein said computer program comprises

a data content transfer component, responsive to user selection of a desired semantic processing operation and user selection of a set of data items within a presented Web page, for:

15 extracting the set of data items from the Web page;

20 providing said set of data items to a respective one of a set of data content processing components running on said Web client which data content processing component is capable of performing the selected processing operation on individual data items extracted from any one of multiple Web pages to generate a result which is dependent on the semantics of the individual data items of the selected set; and

25 providing the generated result to the Web Browser for presentation to the user.

30 12. A Web client data processing system including a set of components according to claim 1.

35 13. A software component for cooperating with a Web Browser, the software component implementing functions to control the operation of a data processing apparatus on which it runs to perform the following steps in response to user selection of a desired data processing operation and user selection of a portion of a Web page presented via the Browser to a user of the data processing apparatus:

40 extracting from the Web page the selected portion of the markup language Web page content;

separating data items from markup tags within the selected portion;

passing the separated data items to a data processing component to perform a processing operation on the individual data items; and

in response to receipt by said software component of a result of said processing operation, generating a modified version of the portion of the markup language Web page content; and

5 providing said modified portion to the Browser to replace the extracted portion of the presented Web page.

14. A software component according to claim 13, wherein the separating of data items from markup tags comprises parsing the selected portion of the markup language content of the Web page to separate data items from 10 markup tags, and parsing the separated data items to remove any extraneous characters.

15. A software component according to claim 13 or claim 14, wherein the generating of a modified version of the portion of the markup language Web page content comprises constructing a markup language separator tag 15 pair for each data item for which results are returned, separately wrapping each of said data items using said tag pair and applying the results to the wrapped data items.

20 16. A software component according to claim 15, wherein data content processing components of said set generate emphasis results including one or more style attributes, and wherein applying said emphasis results to a respective wrapped data item includes adding to the respective tag pair a 25 style parameter including the respective one or more style attributes.

17. A software component according to claim 16, wherein said style attributes include a background colour predefined to correspond to a particular emphasis result.

30 18. A software component according to claim 15, wherein data content processing components of said set generate annotation results including annotation strings, and wherein applying said annotation results to a respective wrapped data item includes adding a segment of dynamic markup language to the respective tag pair which segment activates presentation 35 of a respective annotation string in response to mouse pointer positioning.

40 19. A method for assisting a user of a Web client data processing apparatus with interpretation or analysis of the data content of Web pages presented to the user via a Web Browser, the method including the following steps performed on the Web client data processing apparatus:

45 in response to user selection of a desired data processing operation and user selection of a portion of a Web page presented via the

Browser, extracting from the Web page the selected portion of the markup language Web page content;

separating data items from markup tags within the selected portion;

5

passing the separated data items to a data processing component to perform a processing operation on the individual data items; and

10 in response to receipt by said software component of a result of
said processing operation, generating a modified version of the portion
of the markup language web page content; and

providing said modified portion to the Browser to replace the extracted portion of the presented Web page.

15

20. A computer program product comprising computer readable program code recorded on a computer readable recording medium, the program code including a software component according to claim 13.

30

21. A Web client data processing system including a software component according to Claim 13.

22. A data processing component for interoperating with a software component according to claim 13, including:

25

program code for receiving the separated data items and formatting and/or filtering the data items for processing;

30

program code for performing a processing operation on the formatted and/or filtered individual data items to generate a result; and

program code for returning the result to said software component.

25

23. A data processing component according to claim 22, wherein said generated result includes an annotation string and/or a style attribute for inclusion in a modified version of the portion of markup language Web page content.



Application No: GB 9923173.0
Claims searched: 1-12

30.

Examiner: Mike Davis
Date of search: 27 April 2000

Patents Act 1977
Search Report under Section 17

Databases searched:

UK Patent Office collections, including GB, EP, WO & US patent specifications, in:

UK Cl (Ed.R): G4A (AUDB, APX)

Int Cl (Ed.7): G06F

Other: Online: WPI, EPODOC, JAPIO

Documents considered to be relevant:

Category	Identity of document and relevant passage	Relevant to claims
A	GB 2329492 A (MICROSOFT) eg abstract	-
X	WO 98/18087 A1 (BT) eg abstract, page 19 line 10 to page 25 line 15, and page 32 lines 1-3	1,10,11 at least

X	Document indicating lack of novelty or inventive step	A	Document indicating technological background and/or state of the art.
Y	Document indicating lack of inventive step if combined with one or more other documents of same category.	P	Document published on or after the declared priority date but before the filing date of this invention.
&	Member of the same patent family	E	Patent document published on or after, but with priority date earlier than, the filing date of this application.



Application No: GB 9923173.0
Claims searched: 13-21

Examiner: Mike Davis
Date of search: 14 August 2000

Patents Act 1977
Further Search Report under Section 17

Databases searched:

UK Patent Office collections, including GB, EP, WO & US patent specifications, in:

UK Cl (Ed.R): G4A (AUDB, APX)

Int Cl (Ed.7): G06F

Other: Online: WPI, EPODOC, JAPIO

Documents considered to be relevant:

Category	Identity of document and relevant passage	Relevant to claims
X,E	GB 2346238 A (IBM) whole document	13,19-21 at least
X	WO 98/18087 A1 (BT) eg abstract, page 19 line 10 to page 25 line 15, and page 32 lines 1-3	"

<input checked="" type="checkbox"/> X	Document indicating lack of novelty or inventive step	A	Document indicating technological background and/or state of the art.
<input checked="" type="checkbox"/> Y	Document indicating lack of inventive step if combined with one or more other documents of same category.	P	Document published on or after the declared priority date but before the filing date of this invention.
&	Member of the same patent family	E	Patent document published on or after, but with priority date earlier than, the filing date of this application.